Innovations and Clinical Tools with AI from the National Acoustic Laboratories

Presenter: Jessica Monaghan, PhD; Vicky Zhang, PhD; Matt Croteau, MS

- [Christy] It is my pleasure to welcome a wonderful group of clinicians and researchers from NAL, the National Acoustic Laboratories. Today we have Matt Croteau, Dr. Vicky Zhang, and Dr. Jessica Monaghan to present for us today. If you'd like to learn more about NAL or these presenters today, you can read about their bio on the course registration page, but at this time, I'll hand the mic over to you, Matt.

- Hi. Thank you, Christy. Hi everyone. Good morning, or good afternoon, depending on where you are. I'm going to be starting off with a presentation on our NALguide, but first, just wanna point out, these are our disclosures, so take a second to read over those, and I'll move along in just a few seconds so you have a chance to scan over that. And for today's talk, the learning outcomes for the course are, one, to describe the barriers to widespread hearing screening of children between the newborn period and starting school, and identify potential differences in the speech of children with and without hearing loss. The second one is to describe traditional methods for assessing speech intelligibility performance, and assess the reliability and validity of AI-based methods compared to results obtained from human listeners, and discuss the potential implication for early intervention, and to discuss how to locate, share, and use the NALguide as a resource or a counseling tool.

So that's the one we'll be starting off with is discussing the NALguide. So this is something that you may have seen passed around if you follow anything that we do, but it's a clinical tool basically, that we've recently introduced. And I just wanna present a little bit of background on it, and what it's usefulness is, and how to best use it. So what we'll cover is, first, of course, what is the NALguide, why we actually created this in the first place, how we went about it to make it as valid and useful as possible. And then some examples of how to use the guide. So, of course, you want to know where to find it, what different formats is it available in, put together sort of a made-up, make-believe kind of case use, just to sort of step through it.

And then I'll talk a little bit about what's next, what's coming next after this. So first, what is NALguide? It's a basically a simple booklet that makes hearing aid features easy to understand for people. So it's focused basically on technology, but it connects all these technologies that are in hearing aids to potential benefits. And the whole point of this is to make it as simple and accessible as possible for everyday general public people to understand and to be able to engage and have conversations. So we put it together with a very comprehensive scoping review, and we broke it down into high-level categories that go into further and further detail so people can read a little bit and get the gist of certain technologies and certain benefits that hearing aids have.

And if they have questions or if they have an interest, they can go further and further into the technological details. So the reason why we should have something like this at all, and anyone who's listening, who's a clinician, this is a very easy question to answer and to understand, but it comes down to making information accessible. So one thing we know from research is that 40 to 80% of information given verbally in clinical appointments is forgotten. And that's something that a clinician has seen, you know, a hundred times. And it's not hard to imagine because people are coming in for one appointment, maybe their first ever appointment like this, and there's a million things going on in their heads.

Maybe they've done, you know, like a hearing test that was really distracting and really intense, and then you're starting to tell them all this information. So they tend to forget a lot in clinical appointments. The first fitting appointment is probably even more overwhelming and distracting. So 25 to 50% of information or so, in those appointments is forgotten. And we've also done a review that showed that the printed hearing aid guides and information provided about hearing aids is 69% not rated as suitable, and 31% is rated as adequate. And that comes from everything, from the size and coloring of the text and the print to also just the images and mostly the language that's used. So it's not the most accessible or clear language.

So making information inaccessible in all these different ways results in too much, too little information, or just incomprehensible. It's over people's heads, or they miss one point, and then they don't understand anything that comes after. So information overload or underload is a big problem. Now, where do patients get their information? Lots of places, we all know. They go online, they have marketing materials, they talk to friends, family, doctors, they, of course, see their hearing care professionals, they get marketing, and pamphlets, and data sheets from manufacturers, government programs might have some information. So all of these places are potentially great sources of information, but it can be very confusing and overwhelming, and it can be very inconsistent as well.

So we've recognized this opportunity to try to boil it down and make all of this information very easy to access and understand, at least to a degree where people could understand enough where they could have a conversation and ask the right questions. So really, it's about just demystifying this whole landscape. I think one thing on this slide that I like to point out is that we wanna make sure that this guide is easy to update. So we've done it as future-proof as possible, but we need to keep reviewing it and keep making sure that it identifies emerging technologies. At the moment, we've done a large review, and we've found a few, what we call niche technologies.

So things that one or maybe two manufacturers are doing, but we don't put those off to the side because they're relevant, and maybe next year everybody will be doing it. So we need to sort of keep this up to date. So the process that we went through was, it was born basically out of another project that we were doing, but in the course of doing that project, we identified that this could be a wonderful output of the project. So we had a comprehensive scoping review where we went through all the major manufacturers and we chose a range of their hearing aids from the entry level to the premium. And we started to map and scope all of the features that are listed.

All of this was taken from publicly available data. So we accessed data sheets, marketing materials, pamphlets, flyers, everything we could find about them, we pulled apart all of these features. We then mapped them to what their intended effect was or what their potential benefit was. So you could have something like noise reduction, and you could say, "Well, that could be much, much more comfortable for people if they have noise reduction. It could also be better speech understanding." So we had to sort of start looking at what these features intend to do. We then came up with descriptions for all these categories and these families of features, and then all of these actual individual technologies as well.

And before we finalized anything, and this went on for many months actually, it was quite a process, we did consult with manufacturers to make sure everything was described fairly and that everything was more or less accurate for what their technologies do. And we also consulted with advocacy groups to make sure that the language was appropriate and that the message coming across was very easy to understand and follow. So basically, we had a team of engineers, audiologists, and researchers here working for, you know, a few weeks, several months, really. We extracted over 350 individual hearing aid features. So specifically listed technologies amongst all of these manufacturers. And they were grouped into those families of similar technologies. And then they were assigned an overall benefit or an intended purpose, or sometimes multiple benefits or intended purposes.

The manufacturer group that we consulted with was HAMADAA, which is an Australian hearing aid manufacturing distributor group, and the Deafness Forum Australia as well, which is an advocacy group, and some of the key members we spoke to, but also some additional members that was passed on. And we got a lot of feedback from them. Basically, most of the feedback was about fine-tuning a couple words or just removing jargon and making sure that people could understand it easily. So the end result is a fairly comprehensive guide, but as simple as we could make it. Now you can see here, I've got six sort of main categories, speech understanding, usability, sound quality, connectivity, listening comfort, personalization. This went through a lot of iterations where we had at one point 10 and then we thought maybe we could just make it four or five.

This is where we've ended up at the moment, is just with these six. Every single technological feature that we extracted, we were able to map into one or, in some cases, a few of these categories. And very importantly is that we're not using any manufacturer terms or jargon. Nothing is specific to a manufacturer. It's completely manufacturer-agnostic. It's just about what the technology does. It doesn't matter what it is. And we spent quite a bit of time making sure that it was easy to understand in terms of the reading level, and the jargon, and technical details, and that it was easy to follow. So even visually, we put it together in a way that people could look at it and understand without having to think too hard to understand what was related to what or what page they were on, for example.

So look, there are three main versions of this guide. So we've got an interactive online version, we've got a downloadable printable PDF version, and we do have a downloadable booklet that is similar to the PDF version, but it's formatted to be sent to a professional printer and actually make a leaflet book. So that's available for people, or it's coming soon, I don't think it's on our website yet, but we're just putting that together now. That'll be available, so clinicians or anyone can download that and then have a nice professional book printed. So you can get that at our website at that link. You can also just follow that QR code if you like. Now the Department of Health Hearing Services program in Australia has taken this information and put a simplified version of it on their website as well, so people can access.

It's not a printable version or downloadable version, but people that are accessing that program for information do come across this as well. Now the interactive online version has the greatest level of detail because it sort of unfolds, it's very hard to have all the detail in a booklet without overwhelming people. So the online version has the category, which opens up to feature families with some descriptions about what these groups of features intend to do. And then it goes down into individual features with definitions of what all these things are. And with the drop-down menus, it makes it very clear to navigate. So here's a small kind of screenshot of that. You can see that these are the main categories.

The top one is Speech Understanding. I scroll down, as you can see, so it's cut off the speech understanding part, but then you have Sound Quality, Listening Comfort. It gives you a brief description of what speech understanding means in terms of hearing aids. But then you can click Learn More, and it opens up into all these different feature families. It gives you some key information at the top, and then it says Directional Microphone, explaining what a directional microphone is. And then into some details about the types of directional microphones. These are things that people will come across when they're reading marketing materials, data sheets, reviews for hearing aids. So they can confer to this and say, "Wait, that looks like that's what they were talking about.

And okay, that's what it does." Again, at this level of technical detail at the bottom, we even worked on this to make sure it was at a fairly acceptable reading level and it wasn't using too much jargon. So that's an example of the lower sort of detail, binaural directionality, binaural super directional microphone, things like that. So if people see those terms, they can then just read a simple description about what it is. And then at the bottom, we always have a little bit of a hint which says, "Look out for these words." And those are the types of words that you see in hearing aid literature. Now, for the downloadable and printable PDF version, this has the higher level detail.

So it has the category and this feature families. It doesn't break down into every single individual technological feature and all those details because that would just add several pages, and it makes it a little bit unwieldy. Here's what this one looks like. So here's the speech understanding page. That entire speech understanding category is just all on one page, and it tells you directional microphones, remote wireless microphones, gain and compression, noise reduction, artificial intelligence, and a little bit of a description on how these contribute to speech understanding specifically. Now some of these things are relevant to other categories as well, so there is overlap. Here's an example of the Usability page. So usability, that was actually a very hard category to come up with, but we had lots of loose ends of different features, and we had to work out what's the best way to make, what's the best way to put this into one category together that makes sense.

And even though usability is a little bit vague by itself, if you read the description, you can say, "Oh, I get it. It makes the hearing aids do more things for you." Makes it easier to manage, solves a lot of problems that people have had in the past, for example. So ear-to-ear communication, rechargeability just makes them much more usable and user-friendly. The downloadable booklet that I mentioned, that contains the higher level detail as well, same as the other one. But it has a little bit of background information and a little bit of counseling information as well about how to use the guide and how do you sort of use it as a discussion tool with your friends, your family, your audiologist, whoever.

That's what it looks like. That's sort of a prototype print that we have. That's the downloadable one that you can have professionally printed. And it's got some tick boxes on it. We just thought it was a good idea to have people be able to mark it up in a sensible way so that they can actually remember and recall what they were talking about, either with a friend or with a medical professional. So I've come up with a use case. I can come up with lots and lots of different use cases for this. This one, of course, I've just made up, and I chose to have these sort of elements to it for a specific reason. So this is someone who has been concerned about her hearing.

This is very typical of what you see in the clinic. And she's spoken to lots of people. She has lots of preconceptions about hearing aids and what they may do for her 'cause she knows people who have them. She's asked around, she's been curious about this for a while. She went to a hearing screening event. Somebody said, "Go see an audiologist, you may have hearing loss." And at this event, they had a copy of this NALguide, and she took it home, and she started to leaf through it and noticed a few things. So all she knows about her hearing difficulty is that people are complaining about the TV. Very typical. She's in a book club, for example, so one of her goals might be, I wanna be able to follow conversation meaningfully in my book club.

And she's always asking people to either translate or repeat what her grandchildren are saying to her. She's having trouble hearing that. So she's looking through this book and thinking, "Oh, this might be a little bit difficult, but it's laid out fairly sensibly." And under speech understanding, she starts reading through these different things and might think, "Okay, directional microphones, I've heard about that, I'll circle that one. That's not a bad, it's not a bad one to sort of consider. Noise reduction, I don't know how that works, but I know that noise is a big problem at my book club." Sound quality, she might go through and think, you know, "I don't know what any of that stuff is.

Maybe I can talk about that later." The third one, under the listening comfort, wearing comfort. A lot of friends have told her that hearing aids are really uncomfortable. So she's really concerned about that as something that she needs. Usability, rechargeability. She's got a friend who's got rechargeable hearing aids, for example, and says, "It's actually a lot easier than I thought it would be. It's much better than the other kinds." So she circles that. Connectivity. She just saw that there was a TV streamer she didn't even know that existed so she circled that. That's something she wants to talk about. Personalization, she doesn't need any of that. She doesn't know what that stuff is, but that can all be discussed later.

So just with a glance at very, very simple descriptions, a person is able to sort of pick out things that they think might be important to them without disregarding or forgetting about everything else. The rest of the information is still in the book, but these ones are the ones that she can come in, have already thought about the questions that she wants to ask, and already guide the conversation a little bit. So it saves a bit of time just sort of facilitating the discussion with the audiologist. Now in the clinic, we spend a lot of that time just getting to know people, even during the hearing test, during rapport building, and just getting them settled, you're learning about their hearing difficulties, and you're already preparing this information as a clinician, but for them, they're much more comfortable and much more trusting if they can bring conversation to the table and you're able to reciprocate and say, "That's excellent.

That's exactly what I was gonna talk to you about." So if you look at something like directional microphones, very simple description, "Part of a hearing aid that picks up sound from a given direction, often from in front of you, the microphone can focus on the voices and sounds you want to hear from the direction you want." Just trying to make it as easy as possible for people so they don't have to find out about the physics of how, you know, second and third-order ambisonic microphones work and things like that. So it's all very simple in terms of them, them meaning a patient or anyone who's interested, to pinpoint and remember what they wanna discuss. Just a couple of stock photos that we took just of people using the guide, just showing how it might facilitate in an appointment.

So also, I wanna note that in this review that we did, we also went through assistive listening devices and personal sound amplification devices and made sure that the features available in all of those still map exactly the same. They don't cover all of the features that hearing aid covers, but they don't have features that we didn't cover already. So anyone who's interested in devices like this, you should, more or less, be able to use the guide to have the same discussions. So in terms of what's coming up next and we're doing next. So a few things. We've had a few workshops. My colleague Paola Incerti has been doing some conferences where she's done actual, you know, long sort of workshops with clinicians about this guide, how we may improve it, what we may do in the future, how they feel about it, how they might incorporate it, or how they might use it for their patients.

One thing is that, yes, the printed booklets, a lot of people are interested in us making the printed booklets available already so they can go to medical practices, hearing clinics, pharmacies. We have that PDF version that people can then send off to have printed, but maybe we'll have just printed ones that are available to get directly from us someday. Another thing is that clinicians have generally been expressing an interest in a much more detailed and technologically heavy guide for them. One reason is sometimes there's a certain technology here or there that clinicians don't fully understand, and they're afraid that they can't fully answer a question about particularly new and niche things that come out when manufacturers don't really explain how it works, they just tell you what it might do.

So they kind of want a companion guide that we might put together or start working on that has deeper technological detail. Might go into, for example, that example I had earlier about the microphones. Well, how did two microphones achieve directionality, and why is three better, and why is four better, and how do they become adaptive, and what sort of filters make that work? All of this background information, although not essential to talk to most patients about, it's all really useful because if you're the expert in the chair, you have to be able to field any question, and everyone feels more comfortable if they know 10 times more about what they're talking about than anyone's ever gonna ask them.

So if you know a lot more about the background or the underpinning science of something, you can answer someone's question in five, or six, or seven different ways. You're not stuck. You're not gonna get into a corner where you can only answer it one way, and then you don't know exactly how to be creative after that. So a comprehensive guide with more technological detail might be a good idea, and that would be something that we could keep updated along with the other, the main NALguide as well. So look, we're really just kind of hoping that we can make this useful as possible for people to incorporate in their clinics. I think it's especially useful for pharmacies and general practice because they don't just have, you know, detailed hearing and hearing aid discussions with their patients, but they do need to be able to refer people to information as easily as possible.

So we had quite a large team working on this over the months that we worked on it. The Australian Department of Health and Aged Care helped quite a lot. Deafness Forum Australia helped, and the Hearing Aid Manufacturers and Distributors Association also helped and gave us feedback. In particular, I want to thank Paola Incerti again because she led this project for one thing, but also a lot of this background information and a lot of the feedback that I was able to present to you today, she went and collected for us. So thank you very much. I will move on. I can introduce my colleague, Vicky Zhang, now. She's one of our senior researchers at NAL, and Vicky will be presenting on AI-powered speech intelligibility assessment for children.

All right, Vicky, on to you.

- Sure. Thanks, Matt. Hello everyone. Thank you for joining the webinar today. And so following Matt's topic, I'm very pleased to have this opportunity to talk about a proof-of-concept study that we just finished recently, which we used AI method to assess children's speech intelligibility performance. So let me see if I can control my slide. Oops, I can. Okay, so a little bit background about the speech intelligibility. What is the speech intelligibility, and why it's important? So the speech intelligibility often reflects or refer to how well or clarity of people's speech and that it can be understood by others. So it's an important part of communication, and actually, eligibility is often viewed as benchmark for oral language development because it's kind of a mix of skills of how we hear speech, and how our brain process it, and how we know about the language, and how we pronounce words correctly.

So most typical developing children with normal hearing, they can usually acquire this kinds of skill by the age of four. But how about children with hearing loss? How well their speech can be understood by other listeners, right? Okay, let's have a listen to a sentence spoken by a typical five-years-old children with normal hearing. Let's see if we can play that.

- [Child 1] He is brushing his teeth.

- Okay. And how about this example?

- [Child 1] She is taking a bath.

- Yeah, so the first sentence is, "He is brushing his teeth," and the second one is, "She is taking a bath." Now let's compare these two sentences to those produced by a five-years-old child with hearing loss. Let's try this. Oops. For some reason, I can't play. Can't click the yellow button here. Oops. Hmm. Let's try it again.

- He is.

- Okay, let's try it again.

- Brushing his teeth.

- Good job.

- Okay, let's try it again.

- He is brushing his teeth.

- Good job.

- Okay, so second. Interesting, this always goes to the next slide.

- [Child 2] She is taking a bath.

- All right, so how about the second child, who also has hearing loss?

- [Child 3] He is brushing his teeth. She is taking a bath.

- Okay, as you can tell from these three examples, it can be more challenging to understand speech of children with hearing loss, right? Okay, so let's think about our goal of early intervention. We offer unilateral hearing, universal hearing screening. We offer early diagnosis and early amplification so that children with hearing loss can have necessary support they need. And this helps them hear and develop speech and language just like normal-hearing children, right? The automatic goal that we're trying to do is to help them communicate well, not only with their parents and professionals, but also to make sure everyone they interact with in daily situations can understand what they're saying. Is that right? So the next question is, how we can accurately and efficiently measure the speech intelligibility?

So the method of assessing speech intelligibility depends on a few things, such as what kind of speech materials we choose, how we evaluate it, and expertise of the listeners in listening speech by children with hearing loss. And currently, there are two main methods. One is the transcription method, and the other is scale-based ratings. Traditionally, when we're evaluating children's speech, we have a group of inexperienced listeners to listen to speech, the child's speech and write down what they heard. This is the traditional transcription method. This method is considered to be more accurate compared to the other rating scale method, but the procedure, as you can see, is very time-consuming, and the results can be varied among the different listeners or raters.

Another method to evaluate the speech intelligibility is to use the rating scale. And this method offers a kind of more general way to indicate different levels of speech intelligibility. For example, the seven-point speech intelligibility rating describes the different levels of speech intelligibility performance, which can range from completely intelligible speech to speech that is easily understood by listeners with little or no effort. So each child's speech can be rated by parents, or researchers, or hearing health professionals. And parents can also consider how well their child's speech could understand by others in a casual conversation-like situation and provide a rating. So this method is faster than the traditional transcription method, but it's not as accurate because this method provides more of a overall impression rather than the details accuracy of the transcription method.

And also, it can be varied, depends on who is doing the evaluation. So now I like to introduce our recently completed NAL projects, which we aim to improve the efficiency of the speech intelligibility assessment for young children by using an AI-based method and also to validate whether this method can provide a comparable accuracy results as naive listeners. And also we would like to investigate whether this AI-based method can distinguish between speech intelligibility performance in children with hearing loss and those with normal hearing. So we used the Beginner Intelligibility Test, which called BIT test, BIT sentence test list to measure the speech intelligibility performance. And the test has four sets of 10 short sentences. Each sentence includes two to six words and three to eight syllables.

And we then recruit some children with and without hearing loss. Each child was randomly assigned to one of the four lists and asked to repeat 10 sentences from their assigned list. And we had a total of 2,290 sentences samples recorded by five-years-old native English speaker children without any other disabilities. And among all of these recordings, 850 sentences are from children with normal hearing, and 2,140 sentences are from children with hearing loss. For the hearing loss group, all of the children, they received their cochlear implants or hearing aids before turning three years of age. And we then recruited 69 adult listeners, all of them were between 18 to 40 years old and were native Australian English speakers. They also passed a hearing screening test and had no previous experience interacting with children with hearing loss.

So we call them as naive listeners. And to evaluate the children's speech, we had at least three adults write down exactly what they heard word by word for each sentence. We also use a seven-point signal, sorry, the speech intelligibility rating scale, to assess each child's speech at sentence level and also an overall level after the 10 sentences. Each listener also transcribed the different BIT list for at least the three different children. And each person scored at least one child with normal hearing and one child with hearing loss. And some of the children's speech samples were also evaluated by research participants, research pathologists using the same rating scale. And there are

580 sentences, clean sentences recordings that transcribed and also rated by using the seven-points rating scale by both naive listeners and research speech pathologist.

So this is the rating scale, the seven scale, seven-point rating scale, that we used for this study. Result. First, we selected a few AI transcription models and run them through on children's speech, and then we reviewed the results. Here are two children's speech samples transcribed from different selected AI models and compared them to what naive listeners came up with. The correctly transcribed words are highlighted in green based on the reference list, the first row. For the child's example of Q687, we can see that there were some differences among the four naive listeners when transcribing the same sentence recording. As you can see, out of the four listeners, two heard all of the words correctly compared to the reference list, and the transcriptions from AI models one and two align with the results of these two human beings listeners.

But the other AI models didn't provide the results that were as comparable to those of the humans. And when we took a closer look at the overall accuracy of each AI transcription models by comparing their performance against a reference list, we noted a wide range of accuracy among the 18 selected AI models. And the best-performing model achieved 82.1% accuracy, while the least accuracy was 23% when transcribing the speech recordings of children with normal hearing. And when we applied the same top-performing AI model to transcribe the recordings for children with hearing loss, the accuracy was 71%, which is slightly below the 79% average accuracy that achieved by naive listeners. Now the next question is to assess the agreement of the AI model's performance with that of humans.

So to do this, we used Bland-Altman plots to visualize how closely the AI matches human judgments and if there are any potential bias or systematic differences between the two methods or raters. And on this slide, the average against the difference between the AI model and each of the three listeners, three naive listeners judgments are presented by the three purple Bland-Altman plots. And we also have a comparison between the AI model and the average results from all three naive listeners, which is shown by the green figure. And in each figure, the x-axis represent the average number of the correct words between the AI model and the naive listeners, while the y-axis represents their different score.

And the red dotted line represents the average difference, or the bias. And the dashed lines indicate the upper and lower 95% limits of agreement. When assessing the word-level transcription agreement in speech samples of normal hearing children, all four plots shows the relative symmetry, which, like with the points, distributed both above and below the zero, and the mean bias values on the plots are close to zero to all of them. This indicates very good agreement between the AI and human results in the speech sample of normal hearing children. This agreement falls outside the 95% confidence level in just a few sentences. As you can see, for example, the green figure, out of the 240 sentences, only 11, which is 4.6% for out of the 95% confidence level when we compare the agreement between the AI model and the average results from all three naive listeners.

And there are significant correlations in the number of correct words between the AI method and the average judgments of the naive listeners for the normal hearing group. So overall, these figures demonstrate that there is a strong agreement in the AI's word identification performance across the three individual naive listeners for normal hearing children's speech samples. So about the agreement results for speech samples of children with hearing loss. So similar to the three purple figures represent the agreement's results between the AI and each of the three individual naive listeners judgments, the green figure shows a comparison between the AI model and the average results from all three naive listeners. And just like with the normal hearing children's speech samples, we can see that the mean bias values in all four Bland-Altman plots are also are all close to zero.

And this indicates a very good agreement between the AI and human results in the speech samples of children with hearing loss. However, when we compare it to the normal hearing children's speech samples, we do notice a slightly higher number of sentences falling outside the 95% competence interval level for hearing impaired children. So, for example, about 4.6% of sentences fall out of the 95% competence level when comparing the AI model to the average results from all three naive listeners. While the last slides shows that it was only 4.6% of normal hearing children's speech samples that then fall out of this range. But anyway, the correlation in the number of correct words between the AI methods and the average results from the three naive listeners still remains significant.

So overall, there's still a strong agreement in the AI's word identification performance across the three individual naive listeners for children with hearing aids in their speech samples. So if we consider the AI model as one of the raters and perform statistical analyze to measure how consistent the results among the raters across different groups of children, which include those with normal hearing, hearing loss group, or CI users only, and hearing aid users only, we find the word-level intraclass correlation coefficients, or the ICC results, shows that among naive listeners only, we observed a high level of consistency with the ICC scores very close to one. The normal hearing group shows the highest ICC scores, while the hearing aids group has the lowest scores.

And when we compare the AI model's results to those of individual naive listeners, the ICC results also show a strong agreement with the similar trends as shown in the normal hearing group and the hearing aids group. We noticed that overall internal consistency is somehow lower in the hearing aids group. So this suggests that there is more variation in transcription among listeners, and the AI model for the recordings of children in the hearing aids group which moves. Sorry, it's getting dark in my room. Yeah, so the children, the children's speech, the children with hearing aids group only, their speech is a little bit harder to be understood. So overall, the findings suggest that the AI model is highly consistent with the results of naive listeners.

And this consistency is similar to the internal consistency that we observed in the human listeners across the speech sample from different groups of children, okay? Now let's have a look at how AI and naive listeners performance when assessing the speech intelligibility outcomes among children using different device groups. We use the seven-point speech intelligibility rating this time that's scored by humans to evaluate children's speech intelligibility performance. And on the left side of the figure, we can see that the speech intelligibility ratings provided by naive listeners for children in three groups, right, normal hearing group, cochlear implants group, and hearing aids group. And the results shows that there are a significant difference in the average speech intelligibility rating scores between children with normal hearing and those using hearing aids.

However, there are no significant difference in the SIR ratings between children with CIs and those using hearing aids or children with normal hearing. And we also compare the speech intelligibility performance for these different groups of children using the same SIR rating score that provided by research speech pathologists, and the results showed no significant difference among the three groups of children. Which means, on the other hand, that, yeah, which means that the professionals didn't differentiate the speech intelligibility performance among the groups of children. How about the AI model's performance? The right side of the figure indicates that the overall percentage correct scores provided by the AI model, and the results indicate that children with hearing aids have significantly lower speech intelligibility performance compared to their normal hearing peers or children with cochlear implants.

And children with CIs, cochlear implants, have similar speech intelligibility outcomes with their normal hearing peers. So this means that AI model cannot only distinguish the speech intelligibility performance between the normal hearing group and hearing aids groups as naive listeners can do. But it can also, the AI model can also differentiate the results between the CI and hearing aids group. Here I would like to highlight some key findings which suggest that the selected, the most optimal AI tool or AI model for speech transcription provides an automated and more efficient way to assess speech intelligibility in young children. And this method demonstrates a high level of accuracy and consistency results that align with the word-level transcriptions provided by naive listeners.

And overall, by using this objective approach, these validated AI methods has the potential to revolutionize the assessment of speech intelligibility and perhaps other speech assessments. So for the next step, we are considering perhaps these methods could be integrated into a telehealth platform that clinicians can use as part of the speech assessment, and also it could be further integrated into a user-friendly mobile app to allow our parents or caregivers to track or monitor their children's speech progress at home. Yes, so that's about this proof-of-concept project, and I would like to thanks my colleagues who worked with me in this projects, especially a great thanks to Arun, who's the main person to help with the AI transcriptions and the models, and yeah.

Thank you. So I'll hand over to my colleague Jess for her presentation.

- Thanks, Vicky. So I'm going to be talking about our project using AI to detect hearing loss from children's speech. So the problem we're trying to tackle here, well, we know that untreated hearing loss can have serious and lifelong consequences, but often it's not actually picked up until children start school. So although in Australia and in many other countries around the world, there's newborn hearing screening and in Australia, most children are tested at birth for hearing loss. However, the number of children who develop hearing loss subsequently to the newborn stage, so either it develops subsequently or it's diagnosed at a later age, is just as great or, in fact, greater than those who are picked up at birth.

So I'm sharing some fitting data from Hearing Australia here, and this shows the age in years at first fitting of children in Australia. So you can see that there's a peak around the newborn age, but then there's another peak later on around five and six. So here we can imagine that children are starting school and perhaps struggling there, and this is where hearing losses are being picked up is one element of that. So before children start school and after the newborn hearing screening, there isn't a widespread screening of children either before school age or even at school age, certainly in Australia and also in many other countries. And so in this period before children start school, it's generally assumed that hearing losses will be picked up by parents or early childhood educators.

But actually, what we see in the literature is that parents aren't very good at recognizing hearing loss in their children. So what we see is that only 30% of the time, parents recognize a moderate, mild or moderate hearing loss in their children, and even only 50% of the time for severe or profound losses. So relying on parental suspicion to identify that a child needs a hearing test is not adequate then. But likewise, it's not really practical to test on a wide scale of children before the age of four, since at this age, they generally need specialist expertise and equipment for testing. So to do this across the population would be very expensive and time-consuming. So if we want to ensure that children aren't starting school with undiagnosed hearing losses, we need a new kind of screening approach that can identify hearing loss in a convenient way before children start school.

So this approach that we're looking at here is inspired by the fact that we know speech development is impacted by hearing loss. So much like in Vicky's talk, where she was showing the impact of hearing loss on intelligibility of speech, we also see this at much earlier ages. So even with infants, we can see that language development is affected by hearing loss. So the literature shows that the acquisition of all the phonemes, all of the different speech sounds is delayed in children with hearing loss. But actually, there's a shorter delay for vowels because they have a longer duration and greater energy than for the consonants. So it kind of makes sense that the more audible sounds are, the more exposure children get to them and the more they're likely to produce the sounds.

So as children get older, speech development is a bit more of a complicated process, but at least in the earlier stages, it's really imitating what they're perceiving. So we can see this as well in the fact that the longest delays are seen in the fricatives, so these high-frequency consonants. So that kind of makes sense for children with high-frequency hearing loss that they're not perceiving those sounds and they're not receiving exposure to them. So there's also other kind of patterns, like previous studies have seen that two-year-olds with hearing loss are three times more likely to miss final consonants in their speech than children without hearing loss. And so it's not just a case of a general delay in speech development, there are particular kinds of patterns which make it different, make their speech different to both typically developing children of their own age or children with speech delays for some other cause.

And so that means we might be able to consider that there's a particular kind of signature or pattern that artificial intelligence can be trained to recognize. So that's the kind of task that it's very good for, these kind of pattern matching. So another motivation for this approach, so a reason we might think it could potentially be successful, is that there have been a number of other studies looking at identifying health conditions through speech biomarkers. So particular characteristics of speech that can be indicative of health conditions. So they range from chronic pulmonary disease, Parkinson's, Alzheimer's, depression, and even COVID-19. So people see differences in, for instance, the speech rate, or maybe the voice pitch, or in sort of more language measures as well.

But there's been a number of studies showing very high degrees of accuracy for picking up these kind of health conditions. And there are even screening tools available as mobile apps that are used in workplaces to screen for respiratory conditions and things like COVID. So yeah, it's not an impossible idea that you can identify health conditions through speech. So to test out this concept, we used some speech data that had been collected in the LOCHI study, the longitudinal outcomes in children with hearing impairment study. So that's been ongoing at NAL, and along with other partners, for 16 years now, I think. So these recordings were made when the children were three years old, and they were doing this deep study of phonology.

And this involves showing them pictures, like here, an elephant, girl, lighthouse. And so the child's prompted to identify what's in the picture, and then their speech is transcribed by a speech pathologist, and they're actually able to then identify specific vowels and consonants that they're getting correct or not correct and identify if those are sort of typical for their stage in development. So there's, you know, they're not expected to produce every word perfectly correctly, but there are, you know, there are ratings for what's typical for each age group. So recordings were made during this testing so that the transcripts could be checked for accuracy. And these are the recordings that we use for our speech analysis. So the speech was manually isolated, so the recordings also contain speech from the speech pathologist and from the parents sometimes.

And so we made sure that we isolated just the speech from the child. First of all, I'll show you some differences in the transcripts. So these were produced, the transcripts produced by the speech pathologist and then we can consider which phonemes were being produced correctly. So this is the percentage correct relative to the normal-hearing children. So overall, in black here, they're around 25 percentage points less correct in the consonants that they're producing compared to the normal hearing children. But most interestingly, there's a kind of pattern to the errors. So some of the consonants, like M and N here, are produced reasonably similarly to the normal hearing children. But with sounds like F and V, S and Z here, these fricatives, some of them are approaching 50 percentage points worse than the normal hearing children.

So again, we're seeing these patterns of differences that it's not just sort of worse across the board, but there's specific kind of sounds that aren't being reproduced by the children with hearing loss. So we can see those differences in the transcripts, but obviously, having a speech pathologist first produce those transcripts is not going to be an automatic process. So we want to be able to directly extract the information from the acoustics of the speech. So we do this by first converting the speech recordings into feature vectors. So basically, these are just numbers which capture different aspects of the signal, and the signal's analyzed in short segments, so a bit less than a second each time. And these feature vectors then tell us bits information about, for instance, the pitch or the formants, the shape, the peakiness of the spectrum, its slope, those kind of things.

And also we can see how it changes over time. So there's also evidence that there's difference in the rate of speech for children with hearing loss, and that can be captured also by our networks. So we trained a machine learning model, this deep neural network, and the way it's trained is basically you have many examples of these feature vectors, each corresponding to short segments of speech less than around a second from the children with normal hearing and the children with hearing loss. So for each segment, it makes a guess as to whether it was from a child with normal hearing or with hearing loss. And then we provide these labels as to the correct answer. And so, based on that, it's able to adjust this network, change the parameters each time it's making these guesses to try and make them more correct to get a better accuracy.

And so then, over many, many training examples, eventually it's learns the differences between the speech from the children with hearing loss and normal hearing. And in the end, after training, we find a high accuracy, so around 97% accuracy overall, and that's quite balanced between the children with hearing loss and with normal hearing. So what we see is really good results, especially considering this is quite a small training set in machine learning terms, and actually, this accuracy compares quite favorably to traditional screening techniques with children of this age. And it gives us confidence that this is a potential way to screen children that we can use speech to screen hearing in preschool children and potentially in older age groups.

So presumably, the same kind of approaches would also work with school-aged children. And yeah, so there's an obvious opportunity for a more unobtrusive and convenient screening approach. So if the child needs to just speak, for example, then you could imagine running this analysis using a phone app or a smart speaker, or even a website, anywhere you could upload a speech sample. So we'd come up with this idea, but it's really important, obviously, to check that this is something that people want before we go ahead and try to produce it. So our first step was to interview parents of children in the target age range about, firstly, whether they'd ever had any concerns about speech and hearing with their children and what their experience with the current hearing screening pathway was, and also whether they liked the idea of the proposed solution.

And what we found was that almost all of the parents had had some concerns either with speech and hearing, but mostly it was perceived as a speech problem initially. And they actually found that when they went to see their primary care doctor that they didn't suggest. If the parent presented with concerns about their child's speech, then generally the doctor didn't suggest that they should have a hearing screening. And often, when they would see the speech pathologist, then that wasn't something that came up. So there was, in many cases, quite a delay before they actually had a diagnostic hearing test. They identified that they would want a hearing check before they started school. So in New South Wales, where we perform these interviews, there's a vision screening at four years of age.

And so they sort of questioned why there wasn't, at the same time, a hearing screening. They particularly liked the convenience of it, so not having to make these appointments and go to the doctor. But yes, they were very focused on speech development, and so they said that they would want a feature of the app to be that it's testing, not just for hearing loss but other speech conditions as well, so that was a really great addition. But we also heard from them that they wouldn't just go and seek out this kind of app, they would want to be recommended by some kind of trusted authority, so by their physician, or by daycare workers, or some government information.

And yeah, they thought it was important that it came from a reputable source, so they were happy with the idea of NAL or Hearing Australia producing a solution like that. And so we had some really encouraging feedback from the parents. They said that they thought it was amazing, and they can't believe anyone wouldn't want this. So yeah, we were encouraged by their reaction to the idea. So the next stage for the project is to collect a larger speech database. So we were fortunate enough to get funding from the Australian National Health and Medical Research Council to develop this solution as a mobile app. But for the project, we need to get a large and more diverse speech database.

So we need to make sure that it's, you know, not just trained on children in Sydney, and then it's doesn't work for children in Perth. So we're making sure that this, we're collecting a speech database from children in their homes around Australia, and we're targeting this two and a half to three and a half year age range because that's sort of where we think they're getting to be able to be producing enough speech to make this test feasible, and also where it's still quite difficult to test them in other ways. And so our aim is to collect up to 1000 speech samples from up to 1000 children across Australia, and with this database then we'll be able to train an AI-powered app which will be able to deliver immediate, accurate screening results and then come up with recommendations of what the next stages should be.

Another aspect to this project is that we will also be targeting children with otitis media with effusion. So previously, we've been talking about children with permanent hearing loss, but a big, something that impacts a lot of children in this age group is OME, or glue Ear. And it's particularly common in young children, affecting around 80% of Australian children. And it's even more common in Aboriginal and Torres Strait Islander children, where it can be chronic. So it's quite a public health problem in Australia. So we'd like to see whether... So we are going to target up to 200 children with OME who are undergoing grommet surgery and get them to use the app and see if we can particularly identify OME as either a separate signature or whether we can also, whether we can distinguish it from children with normal hearing.

And yeah, so if this is sensitive enough, then it has the potential to be a valuable tool for ongoing screening and monitoring for children who are at risk of OME. Okay, so to show you how the app works, I'll probably skip to the next slide and then come back for some more details while we show you the video.

- [Narrator] Welcome to the children's speech app trial. Please complete the survey before inviting your child to start the trial. When you're ready for the trial, find a quiet place with minimal distractions and background noise. Now you're ready to begin the Elicit Speech Test app. The app will show pictures one at a time and play a prompt such as, "What is this?" The app will begin to record sound as soon as the picture is shown. When your child says the answer, if it is the same as the word shown on the picture, press the Save and Next button. If your child answers wrong, or if your child doesn't respond after a few seconds, press the Hint button to play a prompt.

There are two hints, so if the first doesn't work well enough, you can try the second hint. During the trial time, please try to avoid letting your child press the buttons. Feel free to encourage your child to make a guess of words if they are not sure. There are 61 words in this test, you can see the progress bar at the top of the screen. After every five words, there will be a cartoon displayed to motivate your child. Please feel free to encourage and praise them yourself too. After the picture test, it will move on to the free speech collection. We would like you to encourage your child in conversation while looking at a picture book or playing with toys for 10 minutes.

If playing with toys, please choose something that is not too noisy. And that's it. The app will automatically upload the recordings and collected data to our secure cloud storage system. Your participation voucher will be emailed to you. Thank you.

- Okay, so as you saw, it displays these pictures, and then it asked, "What is this?" And then hopefully, the child responds straight away. But otherwise, there's a hint. So what we wanted to test in the piloting was that the words we'd selected were appropriate. So we selected words we thought should be in the vocabulary of children in this age range, but we just wanted to check that was generally the case. And also that the pictures were effective at producing the target word as well. So, yeah. So hopefully they say dog here and not golden retriever, as some children might. And also, just in general, that they find it interesting enough and also to see how many words they're able to get through.

So we had 61 for the pilot just to see how far they get before they start to lose interest. And then we asked the parent for feedback about what they thought about these questions.

- [Narrator] Welcome to the.

- And yeah, I'll just show you some more screenshots from the app. So these are kind of pictures, and then we had a reward like that, which seemed to be quite popular with the children. And then this spontaneous speech section at the end. So the idea here is that we, so ideally, you wouldn't have to have a test in the end where you prompted the child. It'd be easier if you don't require the child to behave in any particular way at all, and you can just sort of record speech of them playing. So we wanted to see if that was also possible, so we had this spontaneous speech session at the end. Okay, so overall it was, the children found the app at the right level, they found it interesting.

We tested with range of children with normal hearing, and hearing loss, and different socioeconomic backgrounds. And yeah, in general, the words were familiar, and they found it engaging. What we did find was that the video instructions that I showed you, they were a lot more effective. Originally, we had them reading instructions in the app, and although they would say that they understood the instructions and found them perfectly clear, they would not know what to do with the hint, for example, and they would have to ask the researcher. So that's a good tip in general that people seem to get more information more easily when they see a video rather than just reading instructions. Some people thought the test was too long, so we're gonna cut that down to make sure that it's, that the children don't get bored, and also have the ability to pause the test in the middle.

But yeah, overall, they found it fun, and the rewards seemed to be sufficiently entertaining for them. So just to summarize, the proof of concept shows that it's possible to identify hearing loss in young children from a sample of their speech using artificial intelligence. And this has a potential then to be used as a convenient and inexpensive screening tool. And our next step is to collect a large and diverse speech database to increase the robustness of the tool and reduce any potential biases. We're looking at its potential to detect OME as well as permanent hearing loss. And yeah, so that app should be released early next year, and we'll hopefully soon be able to start training new models.

So there was a big team involved from NAL, and so I'd like to thank all of their efforts and also the funding that we receive to do this work. Okay, thank you, everyone. And it's time for our Q&A now.

- [Christy] Thank you, Dr. Monaghan. We're gonna go ahead and open up the floor for questions and comments for the presenters. We have a few here from early on in the presentation for Matt. Matt, there was a member that asked, "Where would I find the downloadable version of the NALguide?" I'll go ahead and drop that link in for everyone, that way they will have access to it.

- Yep. Yeah, if anyone loses the link, it is just available on our website. So you should be able to just go to our website and go to Projects, and NALguide is listed in there and it's got the interactive version that's in the web page, but the downloadable version will be there as well.

- [Christy] And to caveat off of that, Matt, another member asked, "When might the technical audiologist guide be available?"

- So that's actually something we're just starting to work on now. So I have a feeling it'll be next, mid to late next year, probably. But it is something that we're putting together. So we are sort of reiterating some of the review that we had done. And as part of this next step with the review, we wanna make sure we put together all the technical details. Now, from the previous review that we got the NALguide from, we collected all those technical details, but we just filtered a lot of it out and translated it into simple terms. So we have to go back and then pull the details and find a way to explain, in some cases, some very complex, you know, physics in a way that makes sense, you know, for a clinician.

So it is something we're working on now. I'm planning on having a version to present similar like I did today, maybe by mid-next year. But it's something we're discussing and aware of. So we don't want it to fall by the wayside. So I'll optimistically say 2024. We'll see.

- [Christy] I think it's gonna be a great resource for clinicians. One more question here for you, Matt, from your portion of the presentation. Another member asked, "How often do you intend, does NAL intend to review and update this info? You know, with so many changes going on in the evolution of OTCs and hearing aids every year."

- Yeah, it's kind of a tricky question. We started talking about this way at the beginning of this project, actually. One of the things that we aim to do was to make it as future-proof as possible, so that the way that it stands now, it's generally features that will continue to exist, but they may get more complex or they may get improved on, or they may be presented as a collection of features that are in a family. So you may have a hearing aid that has a certain algorithm or even a certain machine learning or something in it that combines lots of existing features into one sort of output. And those features are all probably represented in the guide already.

However, it's important to, you know, for those six categories that we have, we've put AI in every category with a little bit of a generic description saying that the hearing aids are using artificial intelligence, deep neural networks, machine learning to combine features and to figure out environments, and to answer sonic sort of challenges in the background. And so we can probably expand on that as that becomes a little bit more detailed and becomes more of a talking point for the general public. So we do have to sort of start working on that pretty much right away. Now the other thing is with self-fitting, hearing aids that have self-fitting programs and hearing monitoring programs, and lots of counseling tools built into the devices and the apps, a lot of that's not represented in the guide at the moment anyway.

So that has to sort of be updated pretty soon as well to represent that. Lots of people are curious about that OTC space, and they're very, it's very hard to find the clear answers on where they fit into the whole industry and the whole hearing healthcare process. So something we have to start working on pretty much right away. Now, I alluded earlier about how we're reviewing a lot of that scoping now as it is. So to update the current guide, again, I think that will be happening simultaneously as we build this technical sort of clinician guide. I think it will have to be updated sort of annually. And I think the way that we've planned on doing it, besides setting a, you know, calendar reminder, you know, every November, oh, better update that, we periodically update and refresh our website, and as part of revisiting the NALguide and how it functions on the website, we have to review it and we'll update accordingly.

Things that are done, things that we do like that and present publicly, we don't just do a quick little update. So it still requires a research approach. So we still have a very, we still have a team approach and a research, you know, a very meticulous scoping process. So we do have to plan it, and it will take a bit of time, but I think if we can annually update it to just make sure nothing is new that's not reflected, hopefully once a year that's manageable, but we'll see what happens.

- [Christy] Thank you for your answers, Matt. I have a question here. "In relation to the variability in the intelligibility study with hearing aids in children, what factors do you feel affected the outcome, the fitting, the post-fitting training, and also individual ears?"

- Wow, that's really good question. I think a lot of things might be the factors, like the factors that this person mentioned. Well, as for this particular study, we just want to compare the AI performance versus naive listeners performance. So the transcription, we want to see how the transcription is different in terms of the overall outcomes. I would like to clarify a little bit more on the scoring methods 'cause we use the same tool, the BIT, the beginner intelligibility test tool, which is actually developed by Indiana University, as I remember. And we use the, so for this tool, it has the comprehensive, like, scoring method. So each sentence is scored in terms of the percentage of correctly identified words.

So the score are calculated against the target words, but not at the phoneme levels, like the vows and consonants. So we use the same scoring methods to develop the AI-based algorithm and also compare the results with naive listener's results. So basically, back to the question, the study's focus is not to see what kind of factors that affect the children's speech intelligibility rating rather than we just use the tool. We just want to compare the two on the same speech samples and compare the two methods outcomes.

- [Christy] Thank you, Dr. Zhang. Another question here. "The two traditional methods described in the intro for assessing the SI performance." I think we might need that member to expand on the question. We'll come back to that. Sandra asked, "What is your plan for using your SI research in other countries so that we can use this as a screening tool in the US, here in the States?"

- Oh, that's a really good question. I think from a lot of, like, literatures, the SIR ratings have been kind of widely used everywhere. It should be no problem to be used in US 'cause we actually referred, we find this tool from some papers that I think published by US researchers. I could forward the informations or the papers that we used before. So about the SIR ratings, is kind of a widely used tool well as the traditional tool because it's kind of time-consuming, so it's not that popular used. But for this study, we try to use the AI methods but go through the traditional ways, more accurate way of transcribing children's speech into word level of accuracy to see the overall speech intelligibility.

So I think in terms of, like, the future, like the slide I showed on the comparison between the SIR ratings across different device groups versus the transcription results by AI, actually seems that AI performs better. I didn't provide a plot, but the SIR rating by researchers, research speech pathologists, actually, it shows no difference between the SI performance across groups. And actually, the results shows that, as you remember, for most of the children's speech intelligibility scores, our researchers rated quite high, always, like, a one to three scores, which means they, because they got professional training so they could understand what the child's speech. Whereas for naive listeners, which is kind of a more useful way to evaluate children's speech, like intelligibility or clarity in the daily life for normal, for naive listeners or AI, they couldn't tell the difference.

But as using the AI method, we can actually distinguish the speech intelligibility outcomes across normal hearing CI and hearing aids group of children. Yeah, does that make sense?

- [Christy] Thank you, Dr. Zhang, for elaborating on that question. Well, we're gonna go ahead and, go ahead and close the classroom. I wanna thank each and every one of the presenters today. I know it is quite early where they're at, and we just appreciate your time and your expertise. I know NAL is, you know, world-renowned and just, like, leading in research, and who better to present on AI than the group here at NAL. So I'd like to thank everyone that put on this presentation, and annually, we put this presentation on as an inner partnership with NAL. And if you've missed any of the previous year's content, that's all available to you on the course library at any time, as well as this one.

And we'd like to thank Matt Croteau, Dr. Vicky Zhang, and also Dr. Jessica Monaghan. Thank you so much for your time. We hope you enjoy the presentation.